PROJECT DOCUMENT
HLG Project: Data Architecture

Project Title	Data Architecture Project
Project Manager: Lead Architect:	Carlo Vaccari Dick Woensdregt
Start Date	January 2018
End Date	December 2018

I.   Background

1.   Statistical organisations deal with many different data sources – each with their own set of characteristics. Statistical organisations need to *find, acquire* and *integrate* data from both traditional and new types of data sources in an ever increasing pace and under ever stricter budget constraints, while taking care of *security* and data *ownership.*

2.   The 2017 HLG-MOS Data Architecture project developed the first version of the Common Statistical Data Architecture (CSDA). This Reference Architecture is a template for NSOs in the development of their own Enterprise Data Architectures. The project also described five different use cases and developed a standard template for the documentation of further use cases.

II.   Project objectives

3.   The objectives of this project are:

   a. To complete the development of the Common Statistical Data Architecture, testing the reference architecture defined in 2017 against other use-cases
   b. To apply and validate the Data Architecture against the outcomes from other groups like UN-GWG, Data Integration project and groups working on statistical ontologies.
   c. To provide guidelines to support statistical organisations in using the Common Statistical Data Architecture.

III.   Scope

4.   The project will focus on providing a more robust version of the Common Statistical Data Architecture as a result of validation against a number of use-cases and integration with the outcomes from other related groups. It will also provide guidance on implementing the architecture.

5.   The project will consider the work done by the 2017 HLG MOS Data Integration project and co-ordinate with other initiatives such as UN Global Working Group on Big Data and the ESSnet on Linked Open Data.

IV.   Contents

6.   This project will comprise of the following two work packages:

Work Package 1: Updated version of Common Statistical Data Architecture

7.   This work package will extend and improve the Common Statistical Data Architecture where appropriate. The work package will test the architecture against further use cases, and include further details on:

       a.   the influence of new data sources on DA
       b.   the connections between Metadata systems and DA
       c.   the relationships between DA and semantic data structures

Work Package 2: Data Architecture Guidelines

8.   This work package will improve the ability of NSOs to implement the Common Statistical Data Architecture. It will develop guidelines that will identify recommendations and steps for the introduction of the Common Statistical Data Architecture in statistical organisations. The guidelines will include:

       a.   descriptions of data artefacts such as Statistical Data Dictionaries, to drive the definition of data structures and metadata
       b.   relationship of the Common Statistical Data Architecture with data governance themes like the data-life cycle management
       c.   best practices to ensure data quality and to share technical solutions (like CSPA services and algorithms)
       d.   roadmap for the adoption and implementation of Data Architecture

V.   Resourcing

9.   The table below shows an estimate of the resources and potential costs involved.

Work Package	Resources (estimate)	Source	Other costs
WP-1	5 person months	Volunteer NSOs plus UNECE Secretariat	Possible travel costs for Sprint
WP-2	9 person months	Volunteer NSOs plus UNECE Secretariat	Possible travel costs for Sprint
Project Management	9 person months	A project manager and lead architect working with the UNECE Secretariat. Input from Executive Board and HLG members (in their role as project sponsors)	Travel costs for project events
Total	23 person months		

## VI.	Timeframe

10.	The project will complete the activities described by the end of 2018. All work packages will start simultaneously and run for the full period of the project.

## VII.	Governance

11.	The project sponsor is the HLG. This is the group that has ultimate responsibility for signing off the project deliverables. In practice, this responsibility will be delegated to the Executive Board. A project manager will have day-to-day responsibility for the running of the project, providing regular updates and signaling any issues to the Executive Board as necessary.

12.	Work will be mainly via electronic communication (email, wiki and WebEx). Face-to-face or virtual sprints can be organized as needed. Other meetings and workshops will be used to facilitate additional face-to-face meetings or to generate discussion. The project team will seek guidance from the Modernisation Groups in the areas of their responsibilities.